October 2019

# Sequential Encoding in Visual Working Memory: In the Absence of Structure, Recency Determines Performance

Jeffery Durbin

SEQUENTIAL ENCODING IN VISUAL WORKING MEMORY:

IN THE ABSENCE OF STRUCTURE, RECENCY DETERMINES PERFORMANCE

A Thesis Presented

by

JEFFERY S. DURBIN

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial fulfillment
of the requirements for the degree of

MASTER OF SCIENCE

September 2019

Psychology

SEQUENTIAL ENCODING IN VISUAL WORKING MEMORY:

IN THE ABSENCE OF STRUCTURE, RECENCY DETERMINES PERFORMANCE

A Thesis Presented

by

JEFFERY S. DURBIN

Approved as to style and content by:

________________________________________
David Huber, Chair

________________________________________
Kyle Cave, Member

________________________________________
Jeffrey Starns, Member

________________________________________
David Moorman, Member

________________________________________
Caren Rotello, Department Chair
Psychological and Brain Sciences

ABSTRACT

SEQUENTIAL ENCODING IN VISUAL WORKING MEMORY:

IN THE ABSENCE OF STRUCTURE, RECENCY DETERMINES PERFORMANCE

SEPTEMBER 2019

JEFFERY S. DURBIN, B.S., UNIVERSITY OF KANSAS

M.S., UNIVERSITY OF MASSACHUSETTS, AMHERST

Directed by: Professor David. E. Huber

Most prior investigations of visual working memory (VWM) presented the to-be-remembered items simultaneously in a static configuration (e.g., Luck & Vogel, 1997). However, in everyday situations, such as driving on a busy multilane highway, items (e.g., cars) are presented sequentially and must be retained to support later actions (e.g., knowing if it's safe to change lanes). In a simultaneous presentation, the relative positions of items are apparent but for sequential presentation, relative positions must be inferred in relation to the background structure (e.g., highway lane markings). To examine sequential encoding in VWM, we developed a novel task in which dots were presented slowly, one at a time, with each dot appearing in one of six boxes (Experiment 1), or in invisible boxes within a visible encompassing outer frame (Experiment 2). Experiment 1 found strong recency effects for judgments of color at the end of the sequence but not for the location of dots. In contrast, without dividing lines, Experiment 2 found strong recency effects for both color and location judgments. These results held true for accuracy, reaction time, and an integrated measure of speed and accuracy. We hypothesize that background structure allows the updating of VWM, slotting each new item into that

structure to provide a new configuration that retains both old and new items, whereas in the absence of structure, VWM suffers from severe retroactive interference.

CONTENTS

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

# CHAPTER 1

## SEQUENTIAL ENCODING IN VISUAL WORKING MEMORY: IN THE ABSENCE OF STRUCTURE, RECENCY DETERMINES PERFORMANCE

The visual world surrounding us is rich and vivid, lush with dynamic and complex data that are transferred to the mind via the suite of sensory and cognitive systems available to humans. Imagine an individual engaging in one such dynamic and complex visual scenario: operating a vehicle on a busy highway. The highway contains clear markings to indicate lanes, the median between opposing directions of traffic, and signage to indicate upcoming exits off the highway. As the individual is driving along the highway, they are surrounded by many cars of varying shapes, sizes, and colors – some of them may even have unique markings such as stripes. Due to the fast speeds the individual and the other cars' drivers are traveling, the individual may find it difficult to "keep track" of all of the cars ahead of their own, or cars that lie behind them. So how might this driver keep track of the deluge of visual information surrounding them? Let us consider two factors: first, we may assume that the individual understands and follows basic rules of driving, remaining focused and cognizant of their surroundings. Second, the driver knows that lanes are clearly marked, which is to say that the driver knows where cars should be located. If the individual were suddenly unable to see the road and could only access information relevant to what had occurred within the past few moments, how would they remember which cars were around them, and where along the highway those cars were located? The present study aimed to further investigate how individuals utilize their visual memories to remember temporally dynamic streams of information on the short term, such as the location and identity of objects in their visual surroundings.

Working memory (WM) is a cognitive system that transfers modality-specific sensory information into a storage mechanism where such information is maintained or operated upon, and has been the subject of several models aimed at explicating the behavioral phenomena of recalling information on the short-term (e.g., Atkinson & Shiffrin, 1968; Baddeley, 2000, 2010; Baddeley & Hitch, 1974). While early investigations in the dynamics of storage, maintenance, and retrieval in WM focused almost exclusively on verbal material, recent research has instead aimed at determining these systemic dynamics in the visual (or visuo-spatial) domain (i.e., referred to as VWM or VSWM, respectively; for a review, see McAfoose & Baune, 2009). Luck and Vogel (1997), for example, developed and investigated the canonical change-detection task, which required individuals to study static arrays of visual information (i.e., colored squares randomly placed throughout the display) and report whether a cued item in a test array had changed from the study phase. Beyond this task, many closely-resembling tasks have been developed to measure different aspects of VWM performance: continuous report tasks, for example, require a participant to report a cued object or object feature on a continuous (often circular) domain, wherein memory precision can be measured as the angular difference between the selected stimulus value and the true stimulus value (e.g., reporting “green” on a pseudo-continuous color wheel when the true value was “blue”; Wilkin & Ma, 2004; Zhang & Luck, 2008). These experimental procedures were primarily used in determining the capacity of VWM: diverging from the seminal work of Miller (1956), the capacity of both verbal and visual working memory appears substantially limited to approximately four objects (for a review, see Cowan, 2000). Vogel, Woodman, and Luck (2001) demonstrated that as the capacity limit for visual information was reached, varying complexity of the items had no effect on participants' performance (i.e., accuracy in recalling single-feature or simple conjunctive items did not differentially suffer from a loss of precision), suggesting that encoding and storage for visuospatial information was an item-general mechanism. This was later discounted with similar studies of stimulus complexity gradients in visual change detection tasks, which showed that (1) complexity gradients sharply decrease an individual's reportable capacity to approximately one item as more features are bound together and the visual object becomes more complex (Hardman & Cowan, 2015), and (2) reportable capacity, and thus the retrieval process, also appeared to depend on the similarity space from which stimuli are chosen (Jackson, Linden, Roberts, & Kriegeskorte, 2015).

Critically, these investigations deducing strict capacity limitations in VWM have focused entirely on simultaneously-presented static arrays of stimuli, where all information is present at once to be encoded, as in the case of the change-detection paradigm. Could capacity limitations arise simply from how much an individual can successfully encode during the study phase? What if the visual information was temporal in nature, where visual information was sequentially available and maintaining the visual information required active maintenance over time? In the present study, these questions were motivated by previous investigations using sequential presentations of verbal material – to understand how one might approach the temporality of visual information, we look first to previous investigations in the verbal domain that have explicated the encoding, maintenance, and retrieval of information from WM.

## Mechanisms of Working Memory

**Verbal domain.** Consider first how information is encoded and stored in the verbal domain (for a review, see Baddeley, 2003). Studies of the mechanisms of maintenance in, search through, and retrieval from WM have been well defined for verbal information (such as written letters or numbers), with Sternberg's (1966, 1969) classic short-term memory scanning task. In this task, a memory set of to-be-remembered numbers, letters, or words was sequentially presented and, after a brief delay, either a positive probe (i.e., a test item that was in the memory set) or negative probe (i.e., a novel item) was presented until a participant responded "old" (i.e., was previously seen during the encoding of the memory set) or "new" (i.e., was not previously seen in the memory set) to the test item. Sternberg examined reaction times (RTs) and identified three particular methods in which participants were able to search through their memory to make judgements about the presence of a particular test item: some participants may search in parallel (i.e., compare the test item to each item in memory simultaneously), some participants may search serially through their memory and terminate their search when a match had been identified, and some would necessarily search through every item in memory regardless if there was a match. Indeed, these theories of search have been shown to be robust and reliable, wherein following studies were able to differentiate between these modes of search -- Monsell (1978) capitalized on these classic findings by comparing sequential presentation conditions in verbal WM paradigms which differed in the rate of presentation (i.e., the duration of each item of the memory set) and the length of delay between the memory set and probe presentation. Interestingly, slow presentation rates (i.e., an item presentation rate of 500 ms or greater per item) with a delay of 800-

1200 ms yielded a pattern of RTs suggestive of serial exhaustive search, which was taken to imply a rehearsal mechanism. Clifton and Birembaum (1970) further showed that a serial self-terminating pattern of RTs could be elicited when the delay between study and test was very brief and the participant was placed under speed demands. This finding established a link between encoding and retrieval processes in verbal WM, which was further accentuated by Monsell's (1978) finding that rapid presentation rates with brief delays yielded a parallel search process with little-to-no rehearsal supporting that process.

From these findings, the capacity limit of verbal WM had been exploited by the *set size effect*, which describes the phenomenon where WM performance suffers as more objects are stored and recalled. In the instance of sequential presentations, the capacity was further able to be evaluated through the *recency/primacy effect* (*or serial position effect, more generally*), which connotes the trend that WM performance is subject to the temporal ordering of serially-presented to-be-remembered items during the study phase, with primacy reflecting increased accuracy for the earliest items in serial position and recency reflecting increased accuracy for the last (few) item in serial position. These effects of set size and recency (i.e., how recent an item in the memory set was seen in comparison to a positive probe) served as the foundation of these claims: in the slow-rate case, an effect of set size on RTs was observed such that as set sizes increased, RTs greatly increased. In the fast-rate case, patterns of RT performance were primarily captured by a recency effect: since RTs remained constant across set sizes, differences in RTs were observed based on the recency of a target-to-probe latency such that the more recent a target was seen in the memory set in comparison to earlier items produced faster RTs. The key demonstration from the verbal literature is that a sequential search pattern of reaction time appears indicative of a rehearsal mechanism, wherein reaction times are dependent on varying degrees of set size but independent from the recency of the remembered items. How might these effects translate into the visual domain?

**Visuospatial domain.** Of principal interest to the present study, of course, is how the systems of encoding and maintenance are characterized in the visuospatial domain (that is, non-verbal, purely visual stimuli). Some evidence has been given to suggest a rehearsal or maintenance mechanism for visual information—for example, Wilson and Emmorey (1997) investigated accuracy performance between deaf patients and American Sign Language (ASL) fluent, age and education-matched healthy controls on a task where participants sequentially viewed signed words in ASL that differed in handshape (i.e., identity), hand location (i.e., position around the face), and hand movement. Stimuli were either similar or dissimilar along these dimensions. Participants responded to a memory set of four signs by signing each after a brief delay. Results from this study, interestingly, found dissimilar stimuli to be more accurately remembered with small suppression effects, suggesting some type of "visuospatial" rehearsal mechanism that allowed individuals to generate and maintain representations of the signs over time along various feature dimension. While informative, the proposition of a visuospatial rehearsal mechanism on the basis of memories of ASL signs cannot be generalized as a description of rehearsal in the visuospatial sketchpad because these visual stimuli are inherently linguistic in nature. This means that evidence of rehearsal indicates an ability to identify and "verbally" label items in the memory set, which does not inform how rehearsal occurs for purely visuospatial information that cannot be easily labeled. More recent and explicit studies of visuospatial material that weren't necessarily linguistic in nature have demonstrated that rehearsal may occur even in tasks where conditions are meant to discourage rehearsal, such as when using a simultaneous presentation of items in visual change detection (Cowan et al., 2016; Donkin, Nosofsky, Gold, & Shiffrin, 2013, 2015).

Interestingly, the use of simultaneous presentations of items in visual change detection tasks may restrict the ability to effectively encode, maintain, search, and retrieve from a visual short-term memory. Mance, Becker, and Liu (2012) compared performance between simultaneous and sequential displays in a visual change detection task with presentation rates/times of the stimuli controlled between experiments, showing that performance (accuracy) between the two conditions was equivalent for low set sizes (i.e., the number of items in the study set, say, 2 items), whereas the sequential condition showed significantly better performance for larger set sizes of 3 or 4 items. This finding suggests that in a sequential paradigm, individuals are able to store and retrieve more effectively (presumably from an increased encoding capacity), making it ideally suited to examine how rehearsal might occur in the visuospatial domain as it can demonstrate both set size effects, as well as recency effects such as those that have been well-described in the verbal literature. Alternatively, Kahana and Sekuler (2002) investigated a visual analogue of the Sternberg search paradigm via a sequentially-presented visual STM task wherein participants observed memory sets of 1, 2, 3, or 4 spatial frequency profiles and, after a brief delay, gave a binary “yes” or “no” response to having seen a probe previously. Their results, which only considered accuracy, suggested strong recency effects such that any differences among set sizes were captured by recency that was superimposed on a similarity gradient within the memory set. Similarly, Nosofsky, Little, Donkin, and Fifec (2011) used a color-based sequential presentation task where colors were mechanistically chosen to form a similarity gradient in color perceptual information in order to explicitly extend the Sternberg search paradigm into the visual domain. Both accuracy and RT results suggested a recency gradient that was a function of similarity and set size in the sequential presentation of visual stimuli. Furthermore, trends in both accuracy and RT suggested that rehearsal was attempted by participants for large set sizes (e.g., greater than 2 or 3 items) which created distortions in the similarity effects that were observed (since, presumably, highly similar colors would be difficult to rehearse). In fact, it is explicitly indicated that the manner of encoding, rehearsal, and its relationship to retrieval is a critical point of interest in the study of the visuospatial sketchpad:

> Future research should continue to investigate the issue. Although our experimental methods were intended to discourage complex rehearsal strategies, they probably did not eliminate them completely. Lawful quantitative relationships between memory strength and lag [recency] may be observed under conditions where rehearsal is brought under tight control, thereby leading to still more parsimonious accounts of memory-scanning performance. (p. 304, footnote 11).

While Nosofksy et al. (2011) used a slow presentation rate in their sequential memory paradigm, Nosofsky and Donkin (2016) constructed a sequentially-presented visual change detection task under fast presentation rates (i.e., 150 ms). Results from RT modeling analyses indicated parallel search with a set size effect explained by strong recency effects guiding both the accuracy and RT data, which further did not provide evidence of rehearsal. In a more recent endeavor to explore serial position effects in VWM using the continuous-report paradigm, Kool, Conway, and Turk-Browne (2014) showed that recency and primacy effects strongly indicated misbinding and swapping, which simply means that participants were unable to create meaningful and distinct representations of items across time, resulting in mistakenly reporting nontarget objects or individual object features (Unsworth, Heitz, & Parks, 2008 found similar results). Contrary to this position, Ricker, Spiegel, and Cowan (2014) showed that loss of information across time did not result from temporal distinctiveness (i.e., creating distinct object representations during study, Peteranderl & Oberauer, 2018), but rather from trace decay (i.e., the stimulus signal and the subsequent representation decay, or lose information, over time; Donkin, Nosofsky, Gold, & Shiffrin, 2013).

**The Present Study**

As reviewed above, few studies have examined sequential encoding in visual working memory and those that have done so failed to find clear evidence of an active maintenance or updating process. Nevertheless, everyday experience, such as keeping track of where different cars are in a multilane highway, suggests that people are able to effectively maintain visuospatial information over time in some circumstances. Here we consider the role that structure may play in supporting the maintenance of visuospatial information -- for example, the lines of the multilane highway may play a crucial role, providing the structure to allow updating of a single spatial pattern that can effectively capture a small number of objects in a single entity, thus avoiding the need to cycle through rehearsed objects. Thus, when a new car is encountered on the road, it is slotted into this structure (updating the existing visual memory rather than adding a new memory). The current study provides this structure by presenting a discrete number of boxes in which objects can appear in a sequential display. In contrast to this spatial configuration structure, no such structure is provided for color information (future studies might examine whether a structure is possible for color, such as by providing a rainbow with gaps to indicate possible colors).

If information in VWM decays in the absence of structure, we anticipate strong recency effects for conditions that fail to provide structure. In this case, set size effects (e.g., slower responses and/or lower accuracy as a function of increasing sequence length) should be entirely explained by recency effects (i.e., on-average worse performance for longer sequences will occur because longer sequences contain items at greater study-test lags). In contrast, if structure allows updating of objects into VWM in a representation that minimizes decay, perhaps by virtue of containing just a single entity (e.g., a single spatial configuration), then recency effects should be minimal and any set size effects might instead reflect swapping errors or other mechanisms that have been used to explain capacity limits for studies that use simultaneous displays of multiple objects.

In Experiment 1, participants observed a sequence of colored dots presented one at a time with 500 ms for each presentation to allow time to update the content of VWM. Each dot appeared within one of six possible boxes placed in a row along the midline. Thus, a spatial structure was provided but no color structure was provided. Sequences varied from two to five in length and within a sequence no color repeated and no location repeated. To assess whether spatial information was maintained in a different manner than color information, one group of participants was given a “Yes”/“No” location recognition test at the end of each trial while a different group of participants was given a “Yes”/“No” color recognition test at the end of each trial. To anticipate the results, Experiment 1 revealed stronger set size effects for the color group, as compared to the location group, but these set size effects appeared to be entirely explained by recency effects, which were also stronger for the color group. To assess whether this difference merely reflected inherent differences between color and spatial representations in general, regardless of structure, Experiment 2 repeated the experiment but with the spatial structure removed, revealing strong recency effects for both color and location.

## CHAPTER 2

## EXPERIMENT 1

### Methods

**Participants.** Thirty (*N* = 30) undergraduate psychology students at the University of Massachusetts, Amherst participated in the present study, which was approved by the university's Institutional Review Board. All participants were compensated with course credit. Prior to the experiment, participants self-reported handedness (left = 3) and normal or corrected-to-normal vision (all participants).

**Stimuli.** The present study used the Psychophysics Toolbox in MATLAB to generate and present stimuli on a computer screen (v. 3.0.14, Kleiner, Brainard, & Pelli, 2007). Each stimulus, which occupied a 43.98-pixel circular area circumscribed by a 14-by-14 pixel transparent square, appeared as one of six high-contrast colored dots (RGB values are listed following the color name): red (255, 0, 0), green (0, 255, 0), blue (0, 0, 255), yellow (255, 255, 0), magenta (255, 0, 255), and cyan (0, 255, 255). These dots could appear in any one of a horizontally-linear set of six conjoined cue boxes, which were 20-by-20 pixels in size (area = 400 pixels) and screen-centered, black in color and with line-widths of 1 pixel. Perceptual masks were used to mask each box after a given trial sequence, which appeared simultaneously within each cue box in the form of a 3-pixel line-width 'X.' Thus, all stimuli appeared within a screen-centered 120-by-20 pixel area, which yielded a horizontal 3.62° and vertical .60° subtended visual angle which was well within foveal perception.

**Procedure.** Given the aims of the present study, two conditions within the experiment were created that required participants to recall a single feature dimension of the stimuli: in the first, participants were asked to recognize whether a test probe was located in a spatial location that a studied item occurred (i.e., the "Location" condition); and in the second, the participants were asked to recognize whether a test probe was the same color as a studied item had been during the study phase (i.e., the "Color" condition). These two conditions served as between-subjects independent variable, whereas manipulations of set size and recency were within-subjects independent variables. This design yielded a two-level experimental fixed factor, wherein participants were randomly sorted into the two conditions. Once placed in an experiment, the two within-subjects factors were manipulated: set size yielded four levels (size of 2, 3, 4, or 5) and recency yielded up to five levels (measured as lag, wherein a lag of 1 signified the most recently studied item and 5 being the most distant item). Lag was necessarily dependent on set size, wherein a set size of two could only produce two levels of lag (1 and 2), a set size of three could produce three levels of lag (1, 2, and 3), and so on. Figure 1 (top panel) illustrates the grand experimental design and the differentiation between experimental conditions.

With these manipulations in place, upon arriving to the laboratory for participation each participant first gave informed consent. Following consent, a brief series of questions were given to each participant about handedness and normal vision. The experiment was administered on a 530.23-by-298.45 mm LCD monitor with a refresh rate of 60 Hz and screen resolution of 1920-by-1080 pixels, with monitor configuration settings set to maximum (e.g., contrast, luminance, etc.). Furthermore, the experiment was administered in a dimly-lit room to maximize visual contrast. When the task began, participants first read instructions (which were identical across subexperiments), which simply described that the participant would observe a sequence of colored dots in different locations and their job was to remember this sequence at test. The instructions emphasized that they should attempt to remember items in the sequence as quickly and accurately as possible. Participants were not informed of any of the factorial manipulations. Following instructions, the task began.

Participants first engaged a 28-trial practice session to orient them to their condition's respective version of the task. For any given trial, the trial sequence began with a screen-centered fixation cross for 1000 ms. Following this, the memory set sequence began: while all six cue boxes remained on the screen throughout the set presentation, one colored dot would appear at a time in a randomly chosen (without replacement) location cue for 500 ms. No mask was provided following the presentation of a colored dot. After all dots had been presented (which was dependent on that trial's set size, ranging from 2 to 5 items), a perceptual mask was used at the end of the memory set sequence to prevent iconic persistence for the last item, which appeared in all six cue locations and lasted for 500 ms. Following the mask, a single test item (either a positive or negative probe) was presented: for the Location condition, the test item (a black dot of the same dimensions as the stimuli) was presented either in a previously-filled box or in a box that did not contain a studied item. For the Color condition, the test item (a colored dot of the same size and one of the size colors as the stimuli) was presented in the center of the screen with the cue boxes having been removed from the screen. To accompany the test probe, the text "Did you see this [condition marker] in the sequence you just saw?", where the condition marker was the word "location" or "color" depending on the condition of the participant, appeared above the probe and remained on-screen until the participant responded. Participants were instructed to respond “Yes” using the ‘J’ keyboard key and “No” using the ‘K’ key. Both the choice and reaction time to make this decision were recorded. For the practice trials, feedback would appear after their responses (e.g., “Correct” or “Incorrect”) for 500 ms; feedback was not offered for the main task to avoid practice effects. After the practice trials, participants then completed 420 trials in the main task in which probe type, set size, and lag were counterbalanced and randomized, yielding a total of 15 data points in each combination probe type, set size, and lag per participant. Since negative test items were necessarily not seen in the study phase, lag was not explicitly balanced for these trials – as such, negative trials accounted for 15 * the set size of a given condition (e.g., for a set size of 2 there were 30 negative trials, for a set size of 3 there were 45 negative trials, and so on).

**Results**

The obtained raw data was initially trimmed to remove outliers by log-transforming RT values to normalize the right-skewed RT distribution within each participant for each probe type. Any trial which had an RT greater or less than 2.5 standard deviations around the mean log-RT was discarded. In addition to the trimming, a hard exclusion rule was used to discard any single participant's data from analysis: if the participant had two or fewer correct trials for any combination of set size and lag, their data were excluded from the analyses. One participant's data fit this criterion and was discarded from analysis, yielding unbalanced sample sizes for the two conditions ($N_{\text{Location}} = 15$, $N_{\text{Color}} = 14$). Following the preprocessing of the data, hit rates and false alarm rates for each set size within a participant were distilled to calculate an unbiased measure of performance, namely *d’*. *d’* was calculated by subtracting the normalized false alarm rate from the normalized hit rate, but the process of normalization is sensitive to edge effects of the rates – thus, a 1/2*n* edge-correction (where *n* refers to the number of trials within the cell) was used for rates that equaled 0 or 1 to avoid edge contamination. Beyond this calculation, accuracy rates, mean log-RTs, and mean RTs (untransformed) were calculated for positive trials within each participant for each combination of set size and lag. In addition to these measures, we sought a singular measure that combined RT and accuracy so as to avoid concern that effects seen in one measure or the other might reflect a speed-accuracy tradeoff. To this end, balanced integration scores (BIS) were chosen due to their relative resistance to speed-accuracy tradeoffs within a participant's data, which mitigates the possibility of interpreting both set size and serial position effects as changes in response caution (Liesefeld, Fu, & Zimmer, 2015; Liesefeld & Janczy, 2019). BIS was calculated by first computing the overall mean accuracy rate and mean RT (untransformed) across conditions, and then using the grand means to standardize each participant's accuracy rates and mean RT for each combination of set size and lag. The standardized RT was then subtracted from the standardized accuracy rates, yielding a BIS measure for each combination of set size and lag. Figure 2 shows the mean *d'* between the two conditions and across set sizes, and Figure 3 illustrates the mean accuracy rates, log RTs, and BIS measures across the two conditions for every combination of set size and lag. To analyze these data, two approaches were taken. In the first, we sought an analogue to simultaneous-presentation tasks which can only vary the number of items presented during the study phase – in particular, we examined differences among the dependent variables without worry of the temporality of the presentations and thus only considered differences among set sizes. In the second, we deconfounded the set size variable by further considering the lag between study-stimulus and target presentations.

**Considering Set Size Only.** A 2-way mixed within-subjects ANOVA was conducted on the *d'* values, using Condition as a two-level between-subjects factor, and Set Size as a four-level within-subjects factor. Due to the Set Size factor having greater than two levels, violations of sphericity were accounted for by calculating the Greenhouse-Geisser ε-adjusted *p*-value (reported as $p_c$ where appropriate). Additionally, because the sample sizes of the two conditions were unbalanced, Type III Sum of Squares were used for the ANOVA since it is robust to the unbalanced design, and also because this type of sum of squares is best suited for detecting significant interaction effects. This analysis yielded a significant interaction effect between Condition and Set Size, $F(2.18, 61.02)$, $p_c = .031$, and a main effect of Set Size, $F(2.18, 61.02) = 9.47$, $p_c < .001$. Given the interaction between Condition and Set Size, a post-hoc 1-way within-subjects ANOVA for each condition was conducted, revealing a significant effect of set size for the Color condition, $F(3, 42) = 12.89$, $p < .001$. Observation of the mean *d'* values for each set size indicated a steady decrease in performance as the number of items within a study sequence increased (i.e., a decrease of 0.89 in mean *d'* from set size 2 to 5). No significant effect of Set Size was observed for the Location condition, $F(3, 42) = 2.56$, $p > .05$. These results suggest that participants' performance in the Color condition appeared to suffer as more items were stored in VWM – but given the sequential presentation in the task, the lag between study and test might underlie this set size effect.

**Considering Lag as a Function of Set Size.** To address the sequential nature of the task and how it may influence performance, this second approach sought to incorporate Lag as a within-subject factor. Lag, which was dependent the set size of any given trial, was not fully-crossed with every level of set size – to account for this, a stratified 3-way mixed within-subjects ANOVA scheme was created for each of the three dependent variables: accuracy rates, mean log-RTs, and the BIS measure. The first test (henceforth referred to as the Lag1-2 Test) of the stratification examined differences in the dependent variables across both conditions, all four levels of set size, and only the first two levels of lag (i.e., lag 1, the most recent item, and lag 2, the second-to-most recent). Similarly, the second test (Lag2-3 Test) in the stratification included only set sizes 3 through 5 and lags 2 and 3, and the third model (Lag3-4 Test) included only set sizes 4 and 5 as well as lags 3 and 4. This stratification scheme was designed for two reasons: first, to ensure that the factorial design was balanced; and second, to allow for a close analysis of how performance changed as a function of the temporal ordering of the stimuli. As with the first approach, Condition was treated as a two-level between-subject factor, and both Set Size and Lag as within-subject factors. In the Lag1-2 and Lag2-3 Tests, the Set Size factor contained more than two levels and thus violations of sphericity were accounted for using the same method as before. Finally, since the interactions between the variables were of chief interest, Type III Sum of Squares were used in conducting the ANOVA with similar justification as in the first approach. Full ANOVA tables are provided in Appendix B.

For the Lag1-2 Test, differences in accuracy rates were examined first. The 3-way ANOVA revealed significant effects of Set Size, $F(3, 81) = 5.06$, $p = .0029$, and Lag, $F(1, 27) = 14.65$, $p < .001$. Since no effect of Condition was observed, post-hoc 1-way within-subjects ANOVAs were conducted for each of the independent variables of Set Size and

Lag, collapsed across conditions. The Set Size simple effect was significant, $F(3, 84) = 5.04$, $p = .0029$, and observation of cell means showed decreased performance for set sizes 3 and 4 ($M_{SS3} = .90$, $M_{SS4} = .88$), but not 2 and 5 ($M_{SS2} = .93$, $M_{SS5} = .91$). Regarding Lag, the simple effect was also significant, $F(1, 28) = 13.69$, $p < .001$, and cell means showed a decreased accuracy rate in Lag 2 as compared to Lag 1 ($M_{L1} = .92$, $M_{L2} = .88$).

Following the analysis of accuracy rates, mean log-RTs were examined. The 3-way ANOVA revealed that all main effects were significant (Condition, $F(1, 27) = 6.69$, $p = .015$; Set Size, $F(2.25, 60.87) = 10.48$, $p_c < .001$; Lag, $F(1, 27) = 53.73$, $p < .001$), and all interactions involving Condition were significant (Condition X Set Size, $F(2.25, 60.87) = 8.21$, $p_c < .001$; Condition X Lag, $F(1, 27) = 6.42$, $p = .017$; Condition X Set Size X Lag, $F(3, 81) = 3.81$, $p = .013$). Because the significant interaction between Condition and Lag is of principal interest given that it provides information about the differences between the two conditions as information is temporally stored in VWM, the post-hoc analysis was comprised of a 2 x 2 mixed ANOVA on mean log-RTs using Condition and Lag as the independent variables for each level of the Set Size factor, accounting for the family-wise error rate (FWE) using a Bonferroni-corrected $\alpha = .0125$. For a set size of 2, a significant effect of Condition was found, $F(1, 27) = 10.50$, $p = .003$, and review of the marginal means by Condition indicated that log RT was faster in the Color condition ($M_{Location} = -0.26$, $M_{Color} = -0.51$, $t(55.05) = 4.51$, $p < .001$). For a set size of 3, significant main effects of Condition, $F(1, 27) = 8.75$, $p = .0063$, and Lag, $F(1, 27) = 21.87$, $p < .001$, were observed. Comparison of marginal means by Condition showed that log RT was faster in the Color condition ($M_{Location} = -0.19$, $M_{Color} = -0.43$, $t(55.68) = 3.95$, $p < .001$), and comparison of marginal means by Lag showed that log RT for lag 1 items were faster ($M_{\text{Difference}}$ = -0.11, $t(28) = -4.75$, $p < .001$). The lack of an interaction between Condition and Lag suggests that the two conditions did not differ on their performance between lags. For a set size of 4, no significant effects were found (all *p*s > .0125). Finally, for a set size of 5, a significant main effect of Lag was observed, $F(1, 27) = 20.29$, $p < .001$, as well as an interaction between Condition and Lag, $F(1, 27) = 14.56$, $p < .001$. Comparisons of performance between lag 1 and 2 for each condition revealed faster log-RTs for lag 1 items in the Color condition ($M_{\text{Difference}}$ = -0.15, $t(13) = -4.88$, $p < .001$), but the same did not hold true in the Location condition ($M_{\text{Difference}}$ = -0.015, $t(14) = -0.77$, $p > .05$).

The Lag1-2 Test was applied to the BIS measures as a means of integrating the effects of both accuracy and RT data. The 3-way ANOVA revealed significant main effects of Set Size, $F(3, 81) = 11.05$, $p < .001$, and Lag, $F(1, 27) = 33.14$, $p < .001$, as well as significant interactions between Condition and Set Size, $F(3, 81) = 5.58$, $p = .0016$, and all three independent variables, $F(3, 81) = 3.81$, $p = .013$. Due to the lack of a significant Condition X Lag or Set Size X Lag interaction, post-hoc analyses comprised of a 2 x 4 mixed ANOVA on mean log-RTs using Condition and Set Size as independent variables for each level of the Lag factor, wherein the FWE rate was again controlled by a Bonferroni-corrected $\alpha = .025$. For a lag of 1, a significant main effect of Set Size was observed, $F(3, 81) = 6.10$, $p < .001$, which was driven by differences in BIS between set sizes 2, 3, and 4 (all *p*s < .02). For a lag of 2, a significant main effect of Set Size, $F(3, 81) = 3.75$, $p = .014$, and significant interaction between Condition and Set Size, $F(3, 81) = 8.43$, $p < .001$, were observed. Further pairwise analyses indicated that in the Location condition, set size 3 differed significantly from set size 5 ($M_{\text{Difference}}$ = -0.78, $t(14) = -4.37$,

$p < .001$). In the Color condition, set size 2 differed significantly from set size 5 ($M_{\text{Difference}} = 0.94$, $t(13) = 5.90$, $p < .001$).

In the Lag2-3 Test, we sought to examine difference in the dependent variables of accuracy rates, mean log-RTs, and BIS measures between conditions only for set sizes 3, 4, and 5, and lags 2 and 3. For accuracy rates, the 3-way mixed ANOVA showed that significant main effects of Set Size, $F(2, 54) = 5.45$, $p = .0069$, and Lag, $F(1, 27) = 16.83$, $p < .001$, were observed. Significant interactions between Condition and Set Size, $F(2, 54) = 4.63$, $p = .013$, as well as Condition and Lag, $F(1, 27) = 20.48$, $p < .001$, were also observed. Given the significant interactions with the Condition factor for both within-subjects independent variables (but with no significant main effect of Condition), consider first the Condition X Set Size interaction. A post-hoc 1-way within-subjects ANOVA conducted for the Location condition showed a significant effect of Set Size, $F(2, 28) = 6.22$, $p = .0058$, which was further supported by significant differences between set sizes 3 and 5 ($M_{\text{Difference}} = -0.044$, $t(29) = -3.13$, $p = .0039$) and between set sizes 4 and 5 ($M_{\text{Difference}} = -0.061$, $t(29) = -3.22$, $p = .0031$). For the Color condition, a significant effect of Set Size, $F(2, 26) = 4.38$, $p = .023$, was observed, further supported by significant differences between set sizes 3 and 4 ($M_{\text{Difference}} = 0.083$, $t(27) = 2.84$, $p = .0083$) and between set sizes 3 and 5 ($M_{\text{Difference}} = 0.055$, $t(27) = 2.71$, $p = .011$). To further examine the Condition X Lag interaction, paired-sample *t*-tests showed a significant decrease in accuracy rates from lag 2 to 3 for the Color condition ($M_{\text{Difference}} = 0.10$, $t(41) = 5.79$, $p < .001$) but not for the Location condition ($M_{\text{Difference}} = -0.005$, $t(44) = -0.34$, $p > .05$).

Similarly, the 3-way mixed ANOVA conducted on mean log-RTs showed a significant main effect of Lag, $F(1, 27) = 22.26$, $p < .001$, as well as interactions between Condition and Set Size, $F(2, 54) = 9.21$, $p < .001$, and between Condition and Lag, $F(1, 27) = 20.48$, $p < .001$. Following the same post-hoc analysis as with accuracy rates, 1-way within-subjects ANOVA using Set Size as the independent variable on mean log-RTs showed only a significant effect of Set Size in the Color condition following Bonferroni-correction, $F(2, 26) = 6.53$, $p = .005$, which was comprised of significant differences between set sizes 3 and 4 ($M_{\text{Difference}} = -0.066$, $t(27) = -2.84$, $p = .0085$) and between set sizes 3 and 5 ($M_{\text{Difference}} = -0.088$, $t(27) = -3.31$, $p = .0026$). Considering the Condition X Lag interaction, paired-sample *t*-tests also showed slower log-RTs for the Color condition between lag 2 and 3 ($M_{\text{Difference}} = -0.16$, $t(41) = -6.11$, $p < .001$), but not for the Location condition ($M_{\text{Difference}} = -0.026$, $t(44) = -1.09$, $p > .05$).

The Lag2-3 Test was lastly applied to the BIS measures. The 3-way mixed ANOVA showed significant main effects of Set Size, $F(2, 54) = 4.36$, $p = .017$, and Lag, $F(1, 27) = 24.82$, $p < .001$. Significant interactions between Condition and Set Size, $F(2, 54) = 13.28$, $p < .001$, as well as between Condition and Lag, $F(1, 27) = 20.16$, $p < .001$, were observed. To further understand the Condition X Set Size interaction, a 1-way within-subjects ANOVA showed a significant effect of Set Size for the Location condition, $F(2, 28) = 8.23$, $p = .0015$, with significant increases in BIS between set sizes 3 and 5 ($M_{\text{Difference}} = -0.72$, $t(29) = -4.13$, $p < .001$) and between set sizes 4 and 5 ($M_{\text{Difference}} = -0.61$, $t(29) = -3.88$, $p < .001$). In the Color condition, the 1-way ANOVA showed a significant effect of Set Size, $F(2, 26) = 9.27$, $p < .001$, with significant decreases in BIS between set sizes 3 and 4 ($M_{\text{Difference}} = 0.82$, $t(27) = 4.31$, $p < .001$) and between set sizes 3 and 5 ($M_{\text{Difference}} = 0.71$, $t(27) = 4.37$, $p < .001$). In considering the Condition X Lag interaction, paired-sample *t*-tests showed a significant decrease in BIS between lag 2 and 3 in the Color condition ($M_{\text{Difference}} = 1.35$, $t(41) = 7.52$, $p < .001$), but no decrease in BIS for the Location condition ($M_{\text{Difference}} = 0.07$, $t(44) = 0.49$, $p > .05$).

Finally, the Lag3-4 Test exploited differences in the dependent variables between both conditions, set sizes 4 and 5, and lags 3 and 4. Regarding accuracy rates, the 3-way mixed ANOVA yielded only a significant interaction between Set Size and Lag, $F(1, 27) = 7.59$, $p = .010$. Paired-sample *t*-tests for set size 5 yielded a significant decrease in accuracy between lag 3 and 4 ($M_{\text{Difference}} = 0.048$, $t(28) = 2.64$, $p = .013$), but the same did not hold true for set size 4 ($M_{\text{Difference}} = -0.033$, $t(28) = -1.19$, $p > .05$). Regarding mean log-RTs, no significant differences were observed from the 3-way mixed ANOVA (all *p*s > .05). Finally, in regard to the BIS measure, a significant main effect of Condition, $F(1, 27) = 4.74$, $p = .017$, and an interaction between Set Size and Lag, $F(1, 27) = 6.85$, $p = .014$, were observed. Overall, the Location condition had higher BIS measures than the Color condition ($M_{\text{Location}} = -0.10$, $M_{\text{Color}} = -1.36$). Paired-sample *t*-tests for set size 5 yielded a marginal decrease in BIS ($M_{\text{Difference}} = 0.46$, $t(28) = 2.12$, $p = .043$), which did not survive Bonferroni-correction, and no change in BIS for set size 4 ($M_{\text{Difference}} = -0.23$, $t(28) = -1.01$, $p > .05$).

**Discussion**

Experiment 1 was conducted in order to test whether temporal information significantly impacted the ability to recall visual information. Across the modeling strata, our data pointed toward the conclusion that performance does suffer as the tested item occurred further back in time during study. In a qualitative comparison of the two approaches, the quasi-“simultaneous” approach appeared to suggest that differences in performance were owed largely to set size effects, with performance being worse in the larger set sizes than the smaller as indicated by lower $d'$ values are higher set sizes. This is ultimately in-line with previous research that capacity could be limited by the number of items (given set size effects), but we expand upon this by arguing that temporal information specifically can further reduce the ability to meaningfully represent all visual information that is presented sequentially. Upon accounting for the temporal nature of the present study's task, it is clear (via the Condition-by-Lag interactions) that not only were the set size effects driven by temporal differences, but that the two conditions of location- or color-based probing also mitigated the lag effects. For example, performance in the Color condition appeared to suffer primarily from lag effects (i.e., a decrease in accuracy as well as an increase in RT in both the Lag1-2 and Lag2-3 Tests, which also transferred to the combinatory BIS measure), whereas the Location condition appeared to suffer primarily from set size effects only as indicated by the lack of differences found between lags on accuracy, RT, and BIS within each of the set sizes.

One might interpret this as a fundamental difference in how location information is accessed versus color information: in the experiment, possible locations and possible colors remained constant. The primary difference, however, was that the possible locations were available to participants at all times – the six cue boxes remained visible at all times, suggesting that the need for temporal information was not needed if the location information need not be explicitly maintained. Thus, in order to make memorial judgements about test items, participants were only required to have a vague idea of which locations were used during study and which were not. Similarly, Z. Gao, Q. Gao, Tang, Shui, & Shen (2015) demonstrated that configural information (i.e., Gestalt cues) could be preserved from perception into VWM and used as a determinant for recalling specific information about the stimulus sequence and the consequent configuration of the stimuli. Taking this into consideration, it appears the Location condition was able to access such information – we note, however, that performance did suffer at larger set sizes (indicated by the Condition-by-Set Size interactions), suggesting that even if this configural information was available for making decisions about what was present in memory, the actual traces of *which* locations were used may deteriorate as more of them were used during the sequence irrespective of their particular position in the sequence. In contrast, configural information was not present by design in the Color condition – even though the location information could serve as a cue for which colors were used during study, the particular arrangement of colors was not well-defined or readily available for the participant. Although the most distinct colors from the spectrum were used (i.e., the primary and secondary colors), it appears that the highly distinguishable colors did not form any meaning relationships between each other so the same strategy could not be employed as in the Location condition. Given the strong presence of recency and lag effects, color information could have decayed as more information was stored into VWM – a last-in, first-out process could govern the trace retention information in this condition, which is supported by the continued lag effects even at later lags of 3 and 4.

To test whether this difference between the color and location conditions reflects an inherent difference between color and spatial information in general, or whether it reflected our procedure, which provided a spatial structure but not a color structure, Experiment 2 removed the spatial configuration information. If performance still differentiated by the condition, it would mean that the access of contextual versus identity information are fundamentally different at the time of retrieval. If, however, performance between the conditions is similar (or at least dependent only on set size), this would indicate that the configural information serves as a means of better accessing location information than color information. As such, Experiment 2 eliminated the presence of specific location cues and instead presented stimuli in a blank space – the same locations as if there were boxes, but with no configuration available to preserve the temporal order or the general spatial frequency of stimuli. In addition to this change from Experiment 1, Experiment 2 collected six times as much data for each participant to allow accurate estimation of effects at the participant level, in light of large individual differences for reaction time data. Although the current study did not do so with its focus on the difference between structure (Experiment 1) versus no structure (Experiment 2), this amount of data for each participant should enable the application of sophisticated reaction time models such as the drift-diffusion or linear-ballistic accumulator models (Brown & Heathcote, 2008; Donkin & Nosofsky, 2012; Nosofsky & Donkin, 2016; Ratcliff, 1978), and we note that the data are freely available as posted on the Open Science Framework (https://osf.io/jzuve/?view_only=45d42e69ebc44aa99a40a8c50a6681b5).

CHAPTER 3

EXPERIMENT 2

## Methods

**Participants.** Twelve ($N = 12$) community members (comprised of undergraduate and graduate students at the University of Massachusetts, Amherst, as well as non-academic community volunteers) participated in Experiment 2, which was approved by the University of Massachusetts, Amherst Institutional Review Board. All participants were financially compensated at a rate of $15 USD per hour with a $5 USD attendance bonus. Prior to the experiment, all participants reported right-handedness and normal/corrected-to-normal vision.

**Stimuli.** The stimuli and end-of-sequence masks were identical to that found in Experiment 1, however the colored dots appeared in one large horizontal, 160-by-20 pixel rectangular area outlined in a 1-pixel black line against a medium gray (128, 128, 128) background..

**Procedure.** As with the stimuli, the procedure remained the same for participation in Experiment 2 (Figure 1, bottom panel, illustrates the experimental procedure for a given trial within the task). In contrast, participants completed three independent, 1.5-hour sessions comprised of 90 blocks of 28 trials (i.e., 30 blocks of 28 trials per session yielding 840 trials per session and 2,520 trials total with 90 data points at each level of set size and lag). Participants were barred from scheduling sessions within two days, so practice was provided before the beginning of each session to ensure familiarity with the task.

## Results

Data were preprocessed in the same manner as in Experiment 1. Four subjects were excluded: two were lost due to attrition (i.e., not completing all three sessions of data collection), and two were lost due to poor performance as defined previously (i.e., having less than 2 correct trials for any combination of set size and lag). This yielded a balanced design of $N = 8$ subjects, with four each in the two conditions ($N_{\text{Location}} = N_{\text{Color}} = 4$). *d'* values were calculated for each subject along each level of set size, and the remaining dependent variables (accuracy rates, mean log-RTs, and BIS measures) were calculated for each subject along each combination of set size and lag. Figures 4 and 5 depict the average of the *d'* values and the accuracy, RT, and BIS data, respectively. Additionally, the two approaches to analyzing the data were taken here as in Experiment 1.

**Considering Set Size Only.** A 2 x 4 mixed within-subjects ANOVA on *d'* scores, treating Condition as a between-subjects factor and Set Size as a within-subjects factor, yielded significant main effects of both Condition, $F(1, 6) = 14.04$, $p < .001$, and Set Size, $F(3, 18) = 50.79$, $p = .0095$. A significant interaction between the two factors of Condition and Set Size was also observed, $F(3, 18) = 11.26$, $p < .001$. Generally, the Location condition showed lower *d'* than the Color condition ($M_{\text{Location}} = 1.43$, $M_{\text{Color}} = 3.36$), and the Color condition featured significant decreases in *d'* between set sizes 2 and 3 ($M_{\text{Difference}} = 1.09$, $t(3) = 4.96$, $p = .015$), set sizes 2 and 4 ($M_{\text{Difference}} = 1.54$, $t(3) = 7.68$, $p = .0045$), set sizes 2 and 5 ($M_{\text{Difference}} = 1.97$, $t(3) = 10.09$, $p = .002$), and finally set sizes 3 and 5 ($M_{\text{Difference}} = 0.88$, $t(3) = 14.12$, $p < .001$). Interestingly, no differences were observed within the Location condition across set sizes (all *p*s > .05), suggesting that *d'* remained constant across set sizes. Of course, this approach overlooks the temporal nature of the task, so next we considered the second approach.

**Considering Lag as a Function of Set Size.** As was done in Experiment 1, a stratified 3-way mixed ANOVA approach was taken to investigate differences among the three dependent variables of accuracy rates, mean log-RTs, and BIS measures. This approach utilized the same three tests to parse the uncrossed Set Size and Lag factors, and thus the same three models were used to analyze the current data – the Lag1-2 Test was used to examine differences for all four levels of set size, and only lags 1 and 2; the Lag2-3 Test for set sizes 3, 4, and 5 as well as lags 2 and 3; and the Lag3-4 Test to investigate differences only for set sizes 4 and 5 and lags 3 and 4. Full ANOVA tables for Experiment 2 are provided in Appendix C.

For accuracy rates, the Lag1-2 Test demonstrated significant main effects of Condition, $F(1, 6) = 8.78$, $p =. 025$, and Lag, $F(1, 6) = 10.66$, $p = .017$. Collapsing across set sizes, accuracy rates were higher for the Color condition than the Location condition ($M_{\text{Location}} = 0.82$, $M_{\text{Color}} = 0.94$, $t(52.08) = -5.28$, $p < .001$), and lag 1 featured higher accuracy rates than lag 2 ($M_{\text{Difference}} = 0.074$, $t(31) = 4.87$, $p < .001$) – the lack of interaction between the Set Size or Lag factors with Condition, however, makes it unclear whether these differences were motivated by Condition. Next, the Lag1-2 Test was applied to mean log-RTs, which yielded significant main effects of Set Size, $F(3, 18) = 11.62$, $p < .001$, and Lag, $F(1, 6) = 9.16$, $p = .023$. A significant interaction between Condition and Set Size, $F(3, 18) = 4.02$, $p = .023$, was also observed. Since Condition did not interact with Lag for mean log-RTs, a post-hoc paired-sample *t*-test collapsed across conditions comparing mean log-RTs between lags 1 and 2 showed faster performance on lag 1 than 2 ($M_{\text{Difference}}$ = -0.13, $t(31) = -6.00$, $p < .001$). To investigate the interaction between Condition and Set Size, a 2 x 4 mixed ANOVA on mean log-RTs showed a significant effect of Set Size for the Color condition, $F(3, 9) = 21.40$, $p < .001$, which was supported by slower mean log-RTs from set sizes 2 to 3 ($M_{\text{Difference}}$ = -.0.14, $t(7) = -2.06$, $p = .0034$), set sizes 2 to 4 ($M_{\text{Difference}}$ = -0.17, $t(7) = -3.95$, $p = .0055$), set sizes 2 to 5 ($M_{\text{Difference}}$ = -0.19, $t(7) = -6.12$, $p < .001$), and finally from set sizes 3 to 5 ($M_{\text{Difference}}$ = -0.059, $t(7) = -4.51$, $p = .0027$). The Location condition again featured no differences in mean log-RTs across set sizes (i.e., no significant main effect of Set Size, $p > .05$, and no pairwise differences between set sizes, all $p$s > .05). Finally, the Lag1-2 Test was applied to the BIS measures, wherein main effects of both Set Size, $F(3, 18) = 7.54$, $p = .0018$, and Lag, $F(1, 6) = 10.90$, $p = .016$, emerged. As in the case of mean log-RTs, the lack of interaction between Condition and Lag suggested that performance across lags 1 and 2 did not depend on the Condition factor, which was supported by lower BIS for lag 2 than lag 1 ($M_{\text{Difference}}$ = 1.10, $t(31) = 5.80$, $p < .001$). For the Condition X Set Size interaction, the same post-hoc approach was taken using the 1-way within-subjects ANOVA: in the Color condition, a main effect of Set Size was observed, $F(3, 9) = 40.07$, $p < .001$, but not in the Location condition, $F(3, 9) = 0.41$, $p > .05$. Further investigation within the Color condition showed decreased BIS from set sizes 2 to 3 ($M_{\text{Difference}}$ = 0.60, $t(7) = 5.23$, $p = .0012$) and set sizes 2 to 5 ($M_{\text{Difference}}$ = 1.23, $t(7) = 5.02$, $p = .0015$).

We used the Lag2-3 Test to investigate differences within the dependent variables for set sizes 3, 4, and 5 along lags 2 and 3. For accuracy rates, only a main effect of Set Size was observed, $F(2, 12) = 4.10$, $p = .040$. Observation of the cell means showed set size 3 having a higher accuracy rate than both set sizes 4 and 5 ($M_{\text{SS3}}$ = 0.84, $M_{\text{SS4}}$ = 0.79,

$M_{SS5}$ = 0.79), but Bonferroni-corrected pairwise *t*-tests revealed no differences among these levels (all *p*s > .05). For mean log-RTs, significant main effects were observed for both Set Size, $F(2, 12) = 4.29$, $p = .039$, and Lag, $F(1, 6) = 13.88$, $p = .0097$, as well as a significant interaction between Condition and Set Size, $F(2, 12) = 7.66$, $p = .0071$. Regarding the main effect of Lag, lag 3 was significantly slower than lag 2 across conditions and set sizes ($M_{\text{Difference}} = -0.072$, $t(23) = -3.73$, $p = .001$). Regarding the Condition X Set Size interaction, a post-hoc 1-way within-subjects ANOVA conducted for each condition showed a significant main effect of Set Size for the Color condition, $F(2, 6) = 19.32$, $p = .0024$, but not for the Location condition, $F(2, 6) = 0.44$, $p > .05$. Further analysis in the Color condition revealed significantly slower log-RTs from set size 3 to 5 ($M_{\text{Difference}} = -0.11$, $t(7) = -5.59$, $p < .001$). The Lag2-3 Test was also applied to the BIS measures, which showed significant main effects of Set Size, $F(2, 12) = 7.97$, $p = .0062$, and Lag, $F(1, 6) = 13.73$, $p = .010$, as well as a significant interaction between Condition and Set Size, $F(2, 12) = 7.69$, $p = .007$. Considering the effect of Lag collapsed across conditions and set sizes, lag 2 had higher BIS measures than lag 3 ($M_{\text{Difference}} = 0.77$, $t(23) = 4.35$, $p < .001$). The post-hoc 1-way within-subjects ANOVA conducted for each condition on BIS revealed a significant main effect of Set Size for the Color condition, $F(2, 6) = 14.51$, $p = .005$, but not the Location condition, $F(2, 6) = 0.43$, $p > .05$. Further analysis showed significantly lower BIS in set size 4 compared to 3 ($M_{\text{Difference,}} = 1.21$, $t(7) = 3.82$, $p = .0064$), and 5 compared to 3 ($M_{\text{Difference}} = 1.22$, $t(7) = 10.07$, $p < .001$).

Finally, the Lag3-4 Test was applied to the dependent variables across set sizes 4 and 5 and lags 3 and 4. Coincidentally, the only significant effects to emerge across the three dependent variables was in the BIS measure: both a significant interaction between Condition and Set Size, $F(1, 6) = 10.66$, $p = .019$, and a 3-way interaction involved Condition, Set Size, and Lag, $F(1, 6) = 7.17$, $p = .036$, were observed. To investigate, paired-sample *t*-tests were conducted using Set Size as the independent variable for each combination of Condition and Lag, yielding only a significant decrease in BIS from set size 4 to 5 ($M_{\text{Difference}} = 1.41$, $t(3) = 3.89$, $p = .03$) in the Color condition and for lag 4.

**Discussion**

Experiment 2 was designed to test whether the differences between the Color and Location conditions in Experiment 1 reflected the inclusion of boxes to provide a spatial structure (i.e., a linkage between the relevant position of one object and the next, considering that they were never on the screen at the same time). In contrast to Experiment 1, both the Location and Color conditions produced strong lag effects in Experiment 2 – when increasing from lag 1 to lag 2, and also from lag 2 to lag 3, there were main effects of lag and yet no interaction between lag and condition (i.e., the lag effects were of the same magnitude for both the Color and Location conditions). Thus, the Color condition was largely the same as Experiment 1 whereas the Location condition revealed a recency effect in Experiment 2 but not Experiment 1. From this, we surmise that the recency effect differences between the Color and Location conditions seen in Experiment 1 reflected the display of discrete boxes during sequential encoding. Interestingly, set size effects were more pronounced in the Color condition than the Location condition in Experiment 2, but this might have reflected a floor effect considering overall poor performance in the Location condition. More specifically, there was a large decrease in performance from lag 1 to lag 2 for the Location condition, with little additional change for higher lags. Thus, in stark contrast to the Location condition of Experiment 1, the Location condition in Experiment 2 exhibited an extreme form of recency, as if only the most recent object was in memory.

The removal of the boxes apparently removed structure for updating spatial representations, but what could provide structure for color? Cave and Bichot (1999) argue that previous findings from Downing and Pinker (1985) as well as Zimba and Hughes (1987) concerning the removal of explicit location information can result in attentional resources being spread across all possible locations (the outer frame of the invisible boxes was displayed in Experiment 2, but such a spread might include anything within this frame). Can color information spread across a range of hues? Some evidence suggests that natural linguistic properties as well as conceptual spaces for color inherently cause grouping among minutely-varied color discrimination tasks (Brouwer & Heeger, 2013; Winawer et al., 2007; Witzel & Gegenfurtner, 2013) but it is unclear if a non-linguistic basis for color structure updating exists. In addition, whether it is possible to provide a non-linguistic structure for color may depend on one’s prior experience with particular sets of colors (e.g., a painter with years of experience using a particular pallet might readily keep track of which colors in that pallet were previously presented).

# CHAPTER 4

# GENERAL DISCUSSION

This study examined whether sequentially presented visual information is processed, maintained, and accessed differently than the canonical approach of using static, simultaneous displays. More specifically, we tested whether background structure (i.e., visual information that sets the stage for sequentially presented items) can differentially affect performance by altering the manner in which the information is maintained (or rehearsed) in VWM. In Experiment 1, colored dots were presented slowly, one at a time, with each dot in a unique location and of a unique color. Background structure was provided in the form of dividing lines between possible locations, providing a framework to encode relative positions even though no two dots appeared at the same time. One group of subjects was given a location recognition test at the end of each trial and for this group of subjects, set size effects were modest and recency effects were absent. In contrast, for a group of subjects given a color recognition test at the end of trial, performance was dominated by strong recency effects. We hypothesize that this occurred because there was no background color structure to enable encoding for the relative positions along the spectrum of possible colors. Experiment 2 removed the spatial structure by removing the dividing lines between the boxes and by widening the outer frame such that dots never appeared in the end positions. In this case, there were strong recency effects for both color and location recognition tests.

We hypothesize that background structure affects the manner in which information is loaded into VWM, supporting a form of chunking by updating the previous configuration to also include the currently presented item. If VWM contains just a single configuration with all previously presented items, this might explain higher accuracy, diminished set size effect, faster responses, and the lack of recency effects. In contrast, in the absence of such structure, each new item overwrites previous items, producing lower accuracy, slower responses, and set size effects that are entirely explained by recency (i.e., a longer list of items is a list that contains items that were farther back in the list). Akyurek, Kappelmann, Volkert, and van Rijn (2017) demonstrated that temporal integration of rapidly-presented stimuli into a visual "chunk" facilitated performance and proved less costly in maintaining the individual stimulus information (as indicated by moderated CDA and P3 amplitudes recorded from electroencephalography) – in the present study, we extend this finding to include instances of visual objects presented at slower rates being "chunked" into an abstract structure that facilitates performance and shields from retroactive interference.

These results may shed light on the relatively small number of studies that have used sequential presentations to study VWM. For instance, Kahana and Sekuler (2002), found recency effects at larger set sizes, and this may have occurred owing to the lack of structure for the spatial frequency profiles used in that study (i.e., the relational information between stimuli was not immediately apparent). Finally, the presence of strong recency effects in the Color condition indicates that the lack of an explicit structure for color information may have contributed to the confusion of items in VWM, as in the case of misbinding, swapping, or interference (Kool, Conway, & Turk-Browne, 2014). It is possible that configural information itself could lend to maintenance of visual information – Jones, Farrand, Stuart, and Morris (1995) demonstrated a similar finding such that memory for sequentially-presented spatial information appeared unhampered by temporal delays so long as spatial cuing was involved to facilitate, suggesting one such "rehearsal" mechanism for the maintenance of a purely visual stimulus (or set of stimuli). Similarly, Nassar, Helmers, and Frank (2018) argue that such a visual "chunking" process actively facilitates encoding and maintenance, thus allowing better performance – Mance, Becker, and Liu's (2012) demonstration of better performance at larger set sizes in the sequentially-presented condition may have been a result of being able to encode the location/color conjunctions within a structure that facilitated performance – even in the simultaneous condition, the structure is provided but the structure is never reinforced and must be encoded at the same time as the stimuli themselves (i.e., it is only seen once, not throughout a sequence of item presentations, thus requiring the immediate processing of the stimuli in conjunction with the structure).

This explanation, that structure within the presentation of stimuli (along feature dimensions), is significant even for research which uses static, simultaneous displays. Because items are presented simultaneously, the structure is constructed for the viewer (i.e., all location-related information is present and all secondary, identity-related information is present) at the time of encoding. Thus, the particular structure for a given feature dimension is necessarily *always* a part of the study and test phases and the role of structural or configural information cannot be exploited in such a static, simultaneous presentation of stimuli. Interestingly, this also provides insight into the ongoing debate of whether the process of encoding and maintenance in VWM reflects the spread of perceptual or attentional resources over a set number of "slots" where items are stored (i.e., the "slots" or "discrete-quanta" family of models; Adam, Vogel, & Awh, 2017; Donkin, Nosofsky, Gold, & Shiffrin, 2013; Rouder et al., 2008; Zhang & Luck, 2008), or rather resources are variably assigned to items as they are encoded and maintained, producing gradations in how precise the memory of a particular item is (i.e., the “variable precision” family of models; Bays & Husain, 2008; Ma, Husain, & Bays, 2014; van den Berg, Shin, Chou, George, & Ma, 2012). If internal structure exists within the feature dimension, it may be the case that these models are in fact attempting to describe the role of such structure as it relates to observed set size effects in simultaneously-presented tasks – these effects may reflect imprecision or specific loss-of-information about a given stimulus as more items are presented within the structure of the stimuli.

Of course, the present study is limited by the lack of direct comparison between the two experiments to determine if the presence of discretized spatial information reliably shielded the visual memory from temporal distortion. Using any parametric statistical test within the framework of null-hypothesis significance testing, the assumption of linearity and equal sample sizes are critical to the success of such tests. In the present study, however, the differences both in the number of trials collected per subject as well as the number of participants within each condition prevented a direct comparison of the Location conditions between the two experiments. Furthermore, if any distributional analyses that compared the distribution of accuracy rates, mean log-RTs, or BIS for each combination of set size and lag were conducted, the larger sample size in Experiment 1 would mean a better-specified distribution than Experiment 2, thus preventing a clear conclusion being reached. To address this, a follow-up experiment could be conducted in two ways to directly test this claim: first, a new experiment could be conducted wherein the trials resembled that of Experiment 2 but with fewer trials per participant and more participants in each condition. Secondly, and more succinctly, a new experiment could be conducted in which participants saw some trials with a discretized spatial arrangement (or frame) and some trials without – the within-subject design would grant more power to detect a difference in performance between having structure and not having structure.

Returning to the scenario of driving on a highway, our evidence makes clear that if an individual suddenly was not able to see the road, their memory of the locations of cars (either their own or others') would fare much better than the actual identities of cars. This may be largely due to the fact that roads are created with a specific spatial configuration – different lanes to hold different cars, signs to tell drivers where they are at or where they are going, and exits off of the highway. In contrast, their memory of specific cars' identities (such as color or shape) would be much less salient because that information is not stable – it changes and is not designed with any configural relations at the time of perceiving the cars, and thus they might only remember which cars were last seen or perceived. It seems that our perception and memory of the visual world around us is subject to a myriad of constraints related to the way we process the information, how the information is maintained, and ultimately how we access that information.

# APPENDIX A

# FIGURES

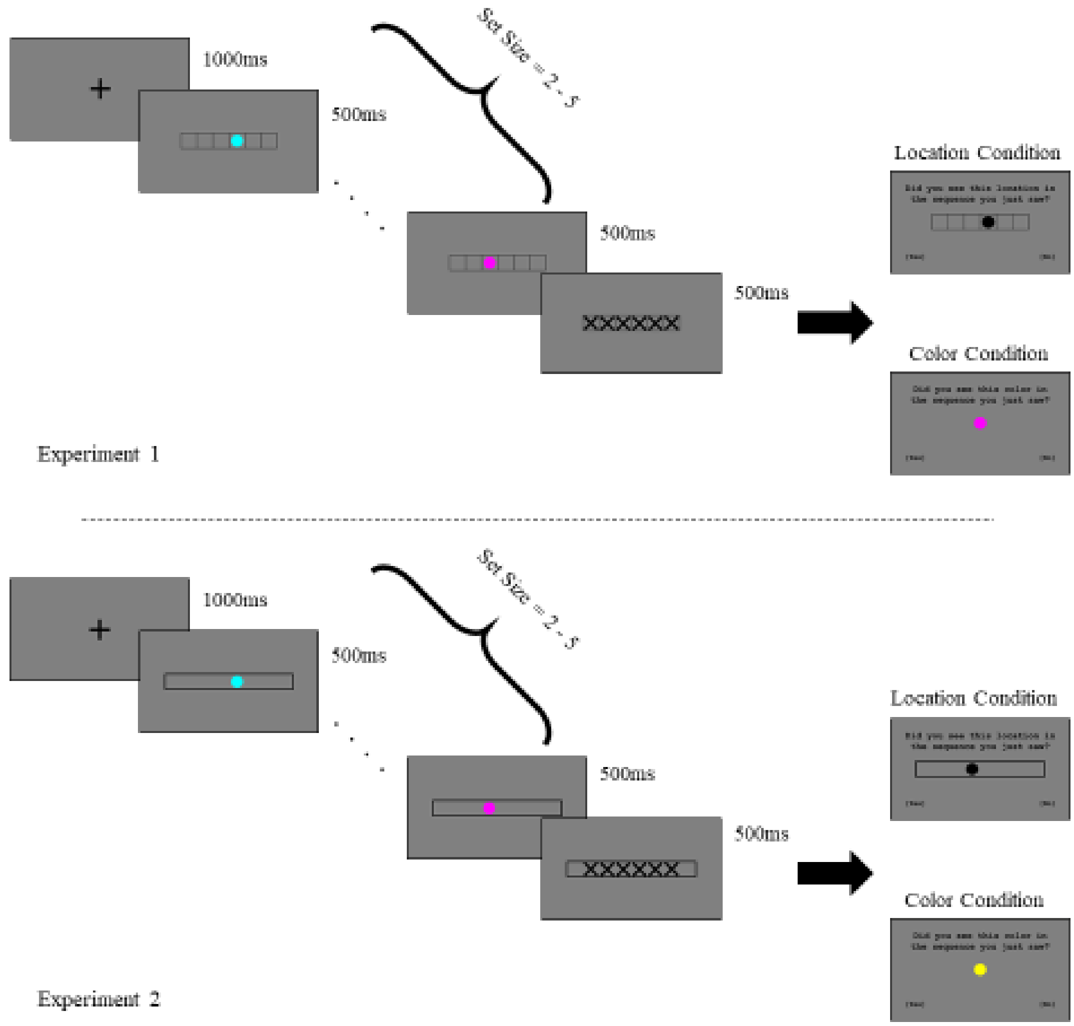


Figure 1: A typical trial.

The typical trial of Experiment 1 (top) and Experiment 2 (bottom). The participant viewed a sequence of colored dots appearing singularly within a spatial cue box (or the large frame as in Experiment 2) – after the mask, participants were either asked to determine if a test item's location (i.e., Location condition) or color (i.e., Color condition) was seen previously during the study phase. The only difference between the two experiments lies in the presence of distinct location cues (Experiment 1) versus a single large frame (Experiment 2).

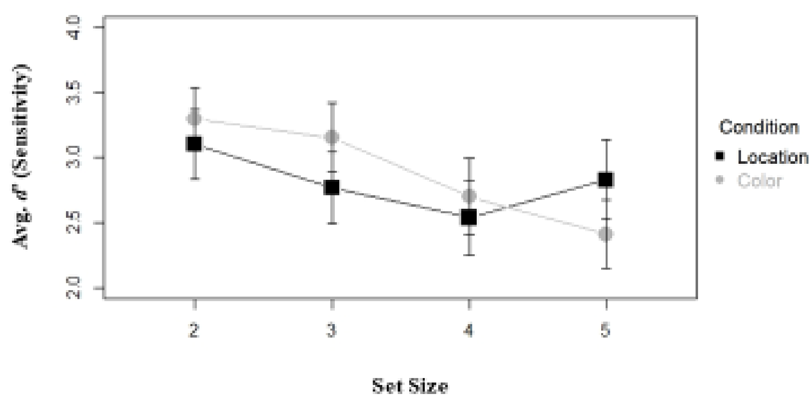


Figure 2: Average *d'* values for experiment 1.

Error bars represent 1 standard error above and below the depicted mean. Both conditions showed relatively high sensitivity, although performance appears to suffer as set size increases.

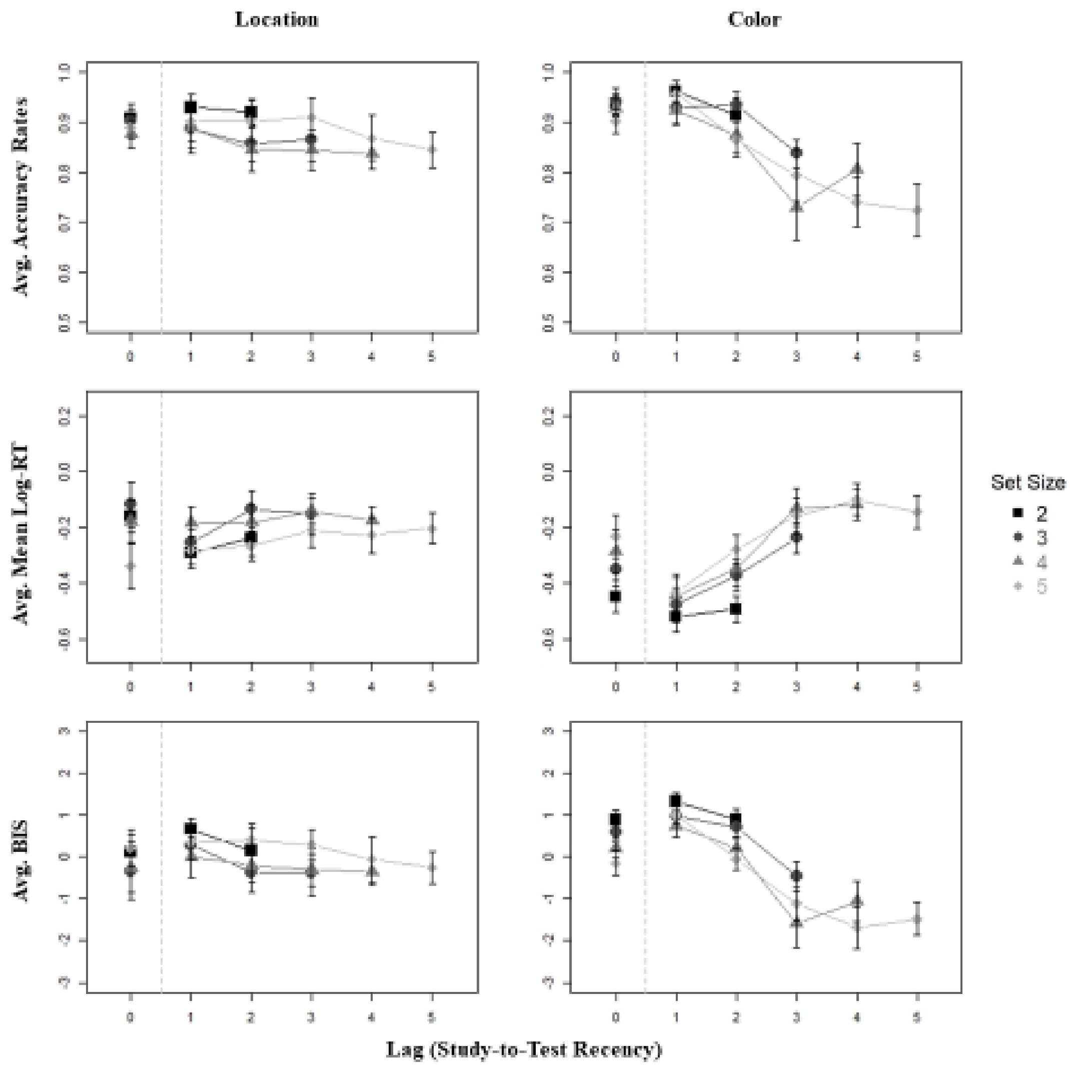


Figure 3: Average performance in experiment 1.

BIS (Balanced Integration Scores) were used as an aggregate measure that combines accuracy and RT data in order to reduce the effects of speed-accuracy tradeoffs and related changes in response caution (Liesefeld, Fu, & Zimmer, 2015; Liesefeld & Janczy, 2019). “Lag 0” refers to the items that were never presented during study (i.e., negative trials). Thus, Lag 0 reflects correct responses to negative trials (i.e., correct rejections) and Lags 1-5 represent correct responses to positive trials (i.e., hits). Error bars represent 1 standard error above and below the depicted mean. As depicted in the figure, the Location condition shows stable accuracy and RT (and consequently the BIS measures) across lag positions, whereas the Color condition shows severe detriments to accuracy and RT as lag positions increase (i.e., the further back in time the item occurred during the study phase, the poorer the performance).

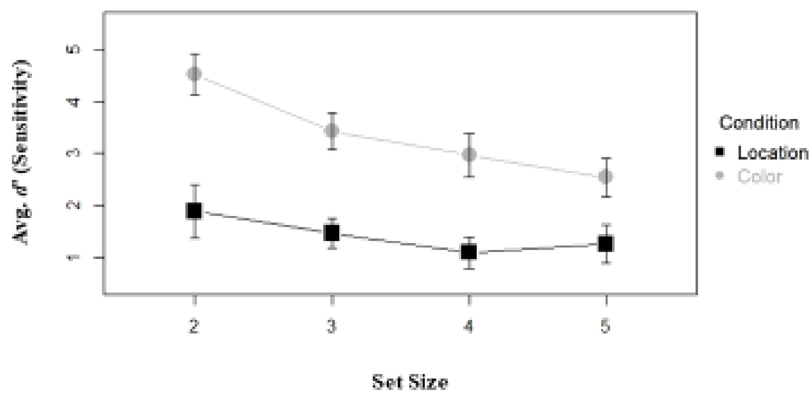


Figure 4: Average *d'* values for experiment 2.

Error bars represent 1 standard error above and below the depicted mean. Generally, the Location condition showed lower sensitivity to determining whether a test item appeared within the study sequence regardless of set size, whereas sensitivity appeared to depend on the set size of the study phase in the Color condition.

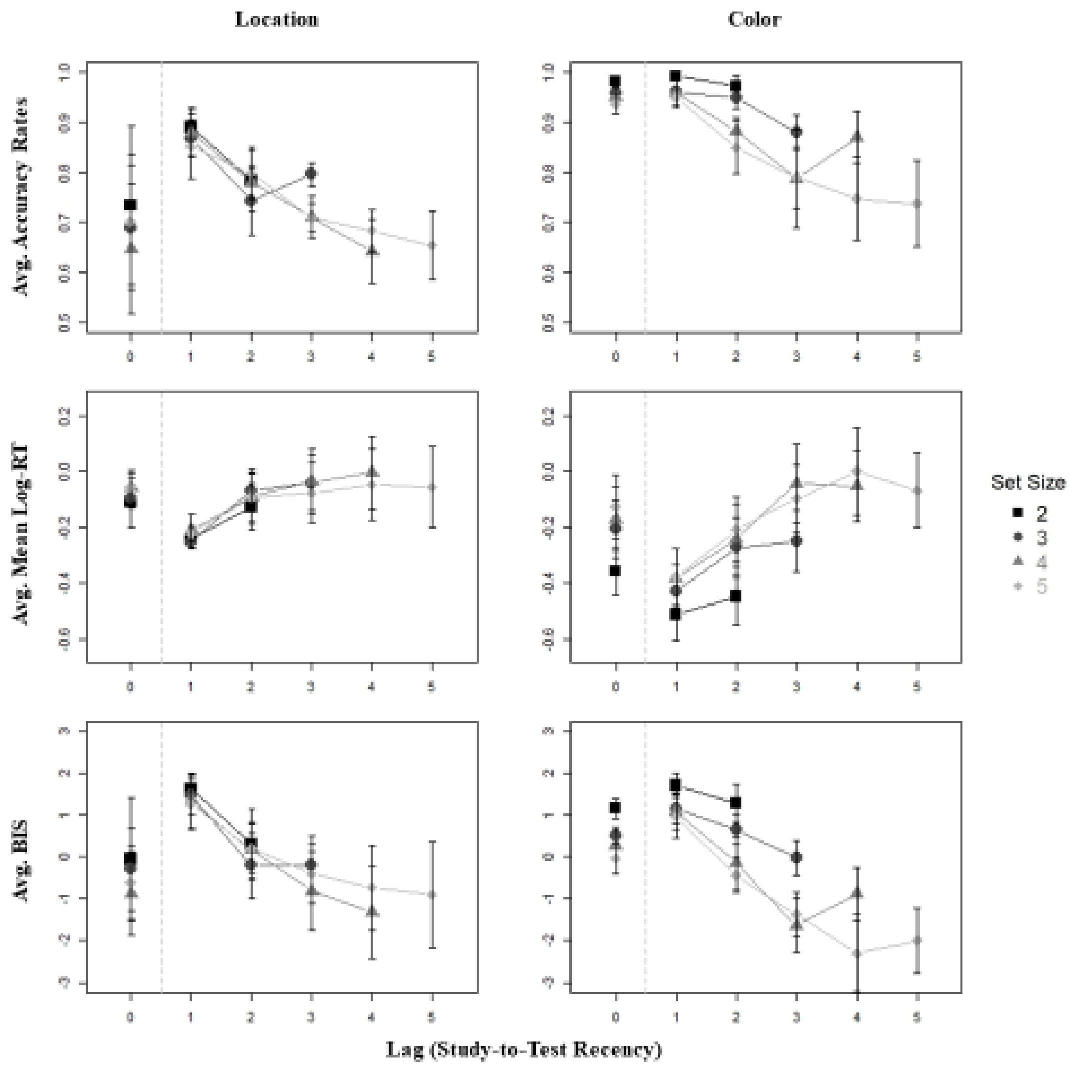


Figure 5: Average performance in experiment 2.

BIS (Balanced Integration Scores) were used as an aggregate measure that combines accuracy and RT data in order to reduce the effects of speed-accuracy tradeoffs and related changes in response caution (Liesefeld, Fu, & Zimmer, 2015; Liesefeld & Janczy, 2019). “Lag 0” refers to the items that were never presented during study (i.e., negative trials). Thus, Lag 0 reflects correct responses to negative trials (i.e., correct rejections) and Lags 1-5 represent correct responses to positive trials (i.e., hits). Error bars represent 1 standard error above and below the depicted mean. Contrary to the findings in Experiment 1, the data depicted here shows lag effects across both conditions (i.e., the conditions become undiscernible such that performance generally decreases according to how distant the item occurred in the past).

# APPENDIX B

## TABULATED ANOVA RESULTS FOR EXPERIMENT 1

| Effect | Numerator *df* | Denominator *df* | *F* | *p* | Significant? |
|---|---|---|---|---|---|
| **Accuracy Rates** | | | | | |
| Condition | 1 | 27 | 0.44 | .51 | |
| Set Size | 3 | 81 | 5.06 | .002 | * |
| Lag | 1 | 27 | 14.65 | < .001 | * |
| Condition * Set Size | 3 | 81 | 1.61 | .19 | |
| Condition * Lag | 1 | 27 | 2.2 | .14 | |
| Set Size * Lag | 3 | 81 | 0.94 | .42 | |
| Condition * Set Size * Lag | 3 | 81 | 2.53 | .062 | |
| **Mean Log-RT** | | | | | |
| Condition | 1 | 27 | 6.69 | .015 | * |
| Set Size | 3 | 81 | 10.48 | < .001 | * |
| Lag | 1 | 27 | 53.73 | < .001 | * |
| Condition * Set Size | 3 | 81 | 8.21 | < .001 | * |
| Condition * Lag | 1 | 27 | 6.42 | .017 | * |
| Set Size * Lag | 3 | 81 | 2.49 | .066 | |
| Condition * Set Size * Lag | 3 | 81 | 3.81 | .013 | * |
| **Balanced Integration Score (BIS)** | | | | | |
| Condition | 1 | 27 | 1.58 | .21 | |
| Set Size | 3 | 81 | 11.05 | < .001 | * |
| Lag | 1 | 27 | 33.14 | < .001 | * |
| Condition * Set Size | 3 | 81 | 5.58 | .0015 | * |
| Condition * Lag | 1 | 27 | 2.63 | .11 | |
| Set Size * Lag | 3 | 81 | 0.12 | .95 | |
| Condition * Set Size * Lag | 3 | 81 | 3.81 | .013 | * |

Table 1: Results of the Lag1-2 Test across dependent variables for experiment 1.

| Effect | Numerator *df* | Denominator *df* | *F* | *p* | Significant? |
|---|---|---|---|---|---|
| **Accuracy Rates** | | | | | |
| Condition | 1 | 27 | 0.36 | .55 | |
| Set Size | 2 | 54 | 5.45 | .0069 | * |
| Lag | 1 | 27 | 16.83 | < .001 | * |
| Condition * Set Size | 2 | 54 | 4.63 | 0.013 | * |
| Condition * Lag | 1 | 27 | 20.48 | < .001 | * |
| Set Size * Lag | 2 | 54 | 1.09 | .34 | |
| Condition * Set Size * Lag | 2 | 54 | 0.65 | .52 | |
| **Mean Log-RT** | | | | | |
| Condition | 1 | 27 | 0.92 | .34 | |
| Set Size | 2 | 54 | 0.98 | .38 | |
| Lag | 1 | 27 | 22.26 | < .001 | * |
| Condition * Set Size | 2 | 54 | 9.21 | < .001 | * |
| Condition * Lag | 1 | 27 | 11.45 | .0022 | * |
| Set Size * Lag | 2 | 54 | 1.45 | .24 | |
| Condition * Set Size * Lag | 2 | 54 | 1.11 | .33 | |
| **Balanced Integration Score (BIS)** | | | | | |
| Condition | 1 | 27 | 0.34 | .56 | |
| Set Size | 2 | 54 | 4.36 | .017 | * |
| Lag | 1 | 27 | 24.82 | < .001 | * |
| Condition * Set Size | 2 | 54 | 13.28 | < .001 | * |
| Condition * Lag | 1 | 27 | 20.16 | < .001 | * |
| Set Size * Lag | 2 | 54 | 1.66 | .20 | |
| Condition * Set Size * Lag | 2 | 54 | 1.32 | .27 | |

Table 2: Results of the Lag2-3 Test across dependent variables for experiment 1.

| Effect | Numerator *df* | Denominator *df* | *F* | *p* | Significant? |
|---|---|---|---|---|---|
| **Accuracy Rates** | | | | | |
| Condition | 1 | 27 | 2.52 | .12 | |
| Set Size | 1 | 27 | 2.04 | .16 | |
| Lag | 1 | 27 | 0.16 | .69 | |
| Condition * Set Size | 1 | 27 | 2.12 | .15 | |
| Condition * Lag | 1 | 27 | 0.94 | .34 | |
| Set Size * Lag | 1 | 27 | 7.59 | .01 | * |
| Condition * Set Size * Lag | 1 | 27 | 2.46 | .12 | |
| **Mean Log-RT** | | | | | |
| Condition | 1 | 27 | 0.73 | .39 | |
| Set Size | 1 | 27 | 1.66 | .20 | |
| Lag | 1 | 27 | 0.04 | .84 | |
| Condition * Set Size | 1 | 27 | 1.19 | .28 | |
| Condition * Lag | 1 | 27 | 1.59 | .21 | |
| Set Size * Lag | 1 | 27 | 0.61 | .44 | |
| Condition * Set Size * Lag | 1 | 27 | 0.14 | .71 | |
| **Balanced Integration Score (BIS)** | | | | | |
| Condition | 1 | 27 | 4.74 | .038 | * |
| Set Size | 1 | 27 | 2.08 | .16 | |
| Lag | 1 | 27 | 0.38 | .54 | |
| Condition * Set Size | 1 | 27 | 3.85 | .06 | |
| Condition * Lag | 1 | 27 | 0.27 | .60 | |
| Set Size * Lag | 1 | 27 | 6.85 | .014 | * |
| Condition * Set Size * Lag | 1 | 27 | 2.1 | .15 | |

Table 3: Results of the Lag3-4 Test across dependent variables for experiment 1.

# APPENDIX C

## TABULATED ANOVA RESULTS FOR EXPERIMENT 2

| Effect | Numerator *df* | Denominator *df* | *F* | *p* | Significant? |
|---|---|---|---|---|---|
| **Accuracy Rates** | | | | | |
| Condition | 1 | 6 | 8.78 | .025 | * |
| Set Size | 3 | 18 | 1.33 | .29 | |
| Lag | 1 | 6 | 10.66 | .017 | * |
| Condition * Set Size | 3 | 18 | 1.19 | .34 | |
| Condition * Lag | 1 | 6 | 0.91 | .37 | |
| Set Size * Lag | 3 | 18 | 0.23 | .87 | |
| Condition * Set Size * Lag | 3 | 18 | 2.18 | .12 | |
| **Mean Log-RT** | | | | | |
| Condition | 1 | 6 | 3.06 | .13 | |
| Set Size | 3 | 18 | 11.62 | < .001 | * |
| Lag | 1 | 6 | 9.16 | .023 | * |
| Condition * Set Size | 3 | 18 | 5.92 | .0053 | * |
| Condition * Lag | 1 | 6 | 0.002 | .96 | |
| Set Size * Lag | 3 | 18 | 1.866 | .17 | |
| Condition * Set Size * Lag | 3 | 18 | 0.7 | .56 | |
| **Balanced Integration Score (BIS)** | | | | | |
| Condition | 1 | 6 | < .001 | .98 | |
| Set Size | 3 | 18 | 7.54 | .0018 | * |
| Lag | 1 | 6 | 0.11 | .016 | * |
| Condition * Set Size | 3 | 18 | 4.02 | .023 | * |
| Condition * Lag | 1 | 6 | 0.38 | .56 | |
| Set Size * Lag | 3 | 18 | 0.43 | .73 | |
| Condition * Set Size * Lag | 3 | 18 | 2.04 | .14 | |

Table 4: Results of the Lag1-2 Test across dependent variables for experiment 2.

| Effect | Numerator *df* | Denominator *df* | *F* | *p* | Significant? |
|---|---|---|---|---|---|
| **Accuracy Rates** | | | | | |
| Condition | 1 | 6 | 3.27 | .12 | |
| Set Size | 2 | 12 | 4.1 | .043 | * |
| Lag | 1 | 6 | 4.72 | .072 | |
| Condition * Set Size | 2 | 12 | 1.68 | .22 | |
| Condition * Lag | 1 | 6 | 0.59 | .47 | |
| Set Size * Lag | 2 | 12 | 1.51 | .25 | |
| Condition * Set Size * Lag | 2 | 12 | 1.36 | .29 | |
| **Mean Log-RT** | | | | | |
| Condition | 1 | 6 | 0.61 | .46 | |
| Set Size | 2 | 12 | 4.29 | .039 | * |
| Lag | 1 | 6 | 13.88 | .0097 | * |
| Condition * Set Size | 2 | 12 | 7.66 | .0071 | * |
| Condition * Lag | 1 | 6 | 4.02 | .091 | |
| Set Size * Lag | 2 | 12 | 4.8 | .029 | ? |
| Condition * Set Size * Lag | 2 | 12 | 2.69 | .10 | |
| **Balanced Integration Score (BIS)** | | | | | |
| Condition | 1 | 6 | 0.12 | .73 | |
| Set Size | 2 | 12 | 7.97 | .0062 | * |
| Lag | 1 | 6 | 13.73 | .01 | * |
| Condition * Set Size | 2 | 12 | 7.69 | .007 | * |
| Condition * Lag | 1 | 6 | 1.5 | .26 | |
| Set Size * Lag | 2 | 12 | 3.27 | .073 | |
| Condition * Set Size * Lag | 2 | 12 | 0.11 | .89 | |

Table 5: Results of the Lag2-3 Test across dependent variables for experiment 2.

| Effect | Numerator *df* | Denominator *df* | *F* | *p* | Significant? |
|---|---|---|---|---|---|
| **Accuracy Rates** | | | | | |
| Condition | 1 | 6 | 1.87 | .22 | |
| Set Size | 1 | 6 | 1.33 | .29 | |
| Lag | 1 | 6 | 0.68 | .44 | |
| Condition * Set Size | 1 | 6 | 5.47 | .057 | |
| Condition * Lag | 1 | 6 | 4.49 | .078 | |
| Set Size * Lag | 1 | 6 | 0.91 | .37 | |
| Condition * Set Size * Lag | 1 | 6 | 3.76 | .10 | |
| **Mean Log-RT** | | | | | |
| Condition | 1 | 6 | 0.0014 | .97 | |
| Set Size | 1 | 6 | 1.34 | .28 | |
| Lag | 1 | 6 | 4.72 | .072 | |
| Condition * Set Size | 1 | 6 | 1.24 | .30 | |
| Condition * Lag | 1 | 6 | 0.16 | .70 | |
| Set Size * Lag | 1 | 6 | 1.7 | .24 | |
| Condition * Set Size * Lag | 1 | 6 | 1.74 | .23 | |
| **Balanced Integration Score (BIS)** | | | | | |
| Condition | 1 | 6 | 0.42 | .54 | |
| Set Size | 1 | 6 | 0.038 | .85 | |
| Lag | 1 | 6 | 1.37 | .28 | |
| Condition * Set Size | 1 | 6 | 10.11 | .019 | * |
| Condition * Lag | 1 | 6 | 0.58 | .47 | |
| Set Size * Lag | 1 | 6 | 4.81 | .07 | |
| Condition * Set Size * Lag | 1 | 6 | 7.17 | .036 | * |

Table 6: Results of the Lag3-4 Test across dependent variables for experiment 2.

BIBLIOGRAPHY

Adam, K. C. S., Vogel, E. K., & Awh, E. (2017). Clear evidence for item limits in visual working memory. *Cognitive Psychology, 97,* 79-97.

Akyurek, E. G., Kappelmann, N., Volkert, M., & van Rijn, H. (2017). What you see is what you remember: Visual chunking by temporal integration enhances working memory. *Journal of Cognitive Neuroscience, 29,* 2025-2036.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence and J. T. Spence (Eds.), *The Psychology of Learning and Motivation: Advances in Research and Theory* (Vol. 2, pp. 89-195). New York: Academic Press.

Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Science, 4,* 419-423.

Baddeley, A. (2003). Working memory and language: an overview. *Journal of Communication Disorders, 36,* 189-208.

Baddeley, A. (2010). Working memory. *Current Biology, 20,* R136-R140.

Baddeley, A & Hitch, G. J. (1974). Working memory. In *The Psychology of Learning and Motivation* (Bower, G. A., ed.), pp. 47-89, New York: Academic Press.

Bays, P. M., & Husain, M. (2008). Dynamic shifts of limited working memory resources in human vision. *Science, 321,* 851-854.

Brouwer, G. J., & Heeger, D. J. (2013). Categorical clustering of the neural representation of color. *Journal of Neuroscience, 33(39),* 15454–15465.

Brown, S. D. & Heathcote, A. (2008). The simplest complete model of choice and response time: Linear ballistic accumulation. *Cognitive Psychology, 57,* 153-178.

Cave, K. R., & Bichot, N. P. (1999). Visuospatial attention: Beyond a spotlight model. *Psychonomic Bulletin & Review, 6 (2),* 204-223.

Clifton, C., & Birenbaum, S. (1970). Effects of serial position and delay of probe in a memory scan task. *Journal of Experimental Psychology, 86,* 69-76.

Cowan, N. (2000). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24,* 87-185.

Cowan, N., Hardman, K., Saults, J. S., Blume, C. L., Clark, K. M., & Sunday, M. A. (2016). Detection of the number of changes in a display in working memory.

*Journal of Experimental Psychology: Learning, Memory, and Cognition, 42,* 169-185.

Donkin, C., & Nosofsky, R. (2012). The structure of short-term memory scanning: An investigation using response time distribution models. *Psychonomic Bulletin Review, 19,* 363-394.

Donkin, C., Nosofsky, R. M., Gold, J. M. & Shiffrin, R. M. (2013). Discrete-slots models of visual working-memory response times. *Psychological Review, 120,* 873-902.

Donkin, C., Nosofsky, R. M., Gold, J. M. & Shiffrin, R. M. (2015). Verbal labeling, gradual decay, and sudden death in visual short-term memory. *Psychonomic Bulletin Review, 22,* 170-178.

Downing, C. J., & Pinker, S. (1985). The spatial structure of visual attention. In M. I. Posner & O. S. Marin (Eds.), *Attention and performance XI: Mechanisms of attention* (pp. 171-187). Hillsdale, NJ: Erlbaum.

Gao, Z., Gao, Q., Tang, N., Shui, R., & Shen, M. (2015). Organization principles in visual working memory: Evidence from sequential stimulus display. *Cognition, 146,* 277-288.

Gilchrist, A. L., & Cowan, N. (2014). A two-stage search of visual working memory: investigating speed in the change-detection paradigm. *Attention, Perception, & Psychophysics, 76,* 2031-2050.

Hardman, K. O., & Cowan, N. (2015). Remembering complex objects in visual working memory: Do capacity limits restrict objects or features? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41,* 325-347.

Jackson, M. C., Linden, D. E. J., Roberts, M. V., Kriegeskorte, N., & Haenschel, C. (2015). Similarity, not complexity, determines visual working memory performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41,* 1884-1892.

Jones, D., Farrand, P., Stuart, G., & Morris, N. (1995). Functional equivalence of verbal and spatial information in serial short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 1008-1018.

Kahana, M. J., & Sekuler, R. (2002). Recognizing spatial patterns: A noisy exemplar approach. *Vision Research, 42,* 2177-2192.

Kleiner, M., Brainard, D., & Pelli, D. (2007). What's new in Psychtoolbox-3? *Perception, 36,* ECVP Abstract Supplement.

Kool, W., Conway, A. R. A., & Turk-Browne, N. B. (2014). Sequential dynamics in visual short-term memory. *Attention, Perception, & Psychophysics, 76,* 1885-1901.

Leisefeld, H. R., Liesefeld, A. M., Muller, H. J., & Rangelov, D. (2017). Fast and careless or careful and slow? Apparent holistic processing in mental rotation is explained by speed-accuracy trade-offs. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41,* 1140-1151.

Leisefeld, H. R., & Janczyk, M. (2019). Combining speed and accuracy to control for speed-accuracy trade-offs. *Behavioral Research Methods, 51,* 40-60.

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature, 390,* 279-281.

Ma, W. J., Husain, M., & Bays, P. M. (2014). Changing concepts of working memory. *Nature Neuroscience, 17,* 347-356.

Mance, I., Becker, M. W., & Liu, T. (2012). Parallel consolidation of simple features into visual short-term memory. *Journal of Experimental Psychology: Human Perception and Performance, 38,* 429-438.

McAfoose, J., & Baune, B. T. (2009). Exploring visual-spatial working memory: A critical review of concepts and models. *Neuropsychological Review, 19,* 130-142.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63,* 81-97.

Monsell, S. (1978). Recency, immediate recognition memory, and reaction time. *Cognitive Psychology, 10,* 465-501.

Nassar, M. R., Helmers, J. C., & Frank, M. J. (2018). Chunking as a rational strategy for lossy data compression in visual working memory. *Psychological Review, 125,* 486-511.

Nosofsky, R. M. & Donkin, C. (2016). Response-time evidence for mixed memory states in a sequential-presentation change-detection task. *Cognitive Psychology, 84,* 31-62.

Nosofsky, R. M., Little, D. R., Donkin, C., & Fifec, M. (2011). Short-term memory scanning viewed as exemplar-based categorization. *Psychological Review, 118,* 280-315.

Peteranderl, S., & Oberauer, K. (2018). Serial recall of colors: Two models of memory for serial order applied to continuous visual stimuli. *Memory & Cognition, 46,* 1-16.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85,* 59-108.

Ricker, T. J., Spiegel, L. R., Cowan, N. (2014). Time-based loss in visual short-term memory is from trace decay, not temporal distinctiveness. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40,* 1510-1523.

Rouder, J. N., Morey, R. D., Cowan, N., Zwilling, C. E., More, C. C., & Pratte, M. S. (2008). An assessment of fixed-capacity models of visual working memory. *Proceedings of the National Academy of Sciences, 105,* 5975-5979.

Sternberg, S. (1966). High-speed scanning in human memory. *Science, 153,* 652-654.

Sternberg, S. (1969). Memory-scanning: Mental processes revealed by reaction-time experiments. *American Scientist, 57,* 421-457.

Unsworth, N., Heitz, R. P., & Parks, N. A. (2008). The importance of temporal distinctiveness for forgetting over the short-term. *Psychological Science, 19,* 1078-1081.

van den Berg, R., Shin, H., Chou, W.-C., George, R., & Ma, W. J. (2012). Variability in encoding precision accounts for visual short-term memory limitations. *Proceedings of the National Academy of Sciences, 109,* 8780-8785.

Vogel, E. K., Woodman, G. F., & Luck, S. J. (2001). Storage of features, conjunctions, and objects in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance, 27,* 92-114.

Wilken, P., & Ma, W. J. (2004). A detection theory account of change detection. *Journal of Vision, 4,* 1120-1135.

Wilson, M., & Emmorey, K. (1997). A visuospatial “phonological loop” in working memory: Evidence from American Sign Language. *Memory & Cognition, 24,* 313-320.

Winawer, J., Witthoft, N., Frank, M. C., Wu, L., Wade, A. R., & Boroditsky, L. (2007). Russian blues reveal effects of language on color discrimination. *Proceedings of the National Academy of Sciences, 104,* 7780-7785.

Witzel, C., & Gegenfurtner, K. R. (2013). Categorical sensitivity to color differences. *Journal of Vision, 13*, 1–33.

Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature, 453,* 233-235.

Zimba, L. D., & Hughes, H. C. (1987). Distractor-target interactions during directed visual attention. *Spatial Vision, 2,* 117-149.